A SURVEY OF PARTIAL DIFFERENTIAL EQUATIONS METHODS IN WEAK KAM THEORY

Lawrence C. Evans *
Department of Mathematics
University of California, Berkeley

CONTENTS

*Supported in part by NSF Grant DMS-0070480 and by the Miller Institute for Basic Research in Science, UC Berkeley

1. Introduction.

This paper is an expanded version of the Courant Lectures I gave at NYU during March, 2002, and provides a idiosyncratic overview of some partial differential equations methods recently developed for "weak KAM theory". I am pushing here the viewpoint that useful information can be extracted from (i) examining two coupled PDE, a generalized "eikonal" equation and a related "continuity" equation, and (ii) exploiting elementary convexity arguments.

I adopt throughout an expository, heuristic style, with the particular aim of emphasizing these two principles: Consult the original papers for precise assertions and full proofs. I discuss mostly my own work (much joint with D. Gomes) and leave out details about the discoveries of Mather, Fathi, and many others. This omission is entirely due to my lack of expertise. While it is currently not so certain that really new dynamical information can be extracted from these arguments, this PDE approach seems to me interesting, at the very least pouring old wine into new bottles.

I thank everyone at the Courant Institute for their kindness and hospitality during my visit. My thanks also to the referee for his/her comments, all of which I have carefully considered, if not fully implemented.

1.1 Lagrangian and Hamiltonian dynamics. We start with a quick and absolutely low-tech review of Lagrangian and Hamiltonian dynamics.

Lagrangian. We begin with a smooth *Lagrangian* $L : \mathbb{R}^n \times \mathbb{T}^n \to \mathbb{R}$, $L = L(v, x)$, where v in $\mathbb{R}^n$ denotes velocity*, and x in $\mathbb{T}^n$, the flat torus in n dimensions, denotes position. Our primary hypotheses will be that

$$\text{the mapping } x \mapsto L(v,x) \text{ is } \mathbb{T}^n \text{ periodic} \tag{1.1}$$

for each $v \in \mathbb{R}^n$, and

$$\text{the mapping } v \mapsto L(v,x) \text{ is uniformly convex} \tag{1.2}$$

for each $x \in \mathbb{T}^n$.

Given a time $T > 0$ and a Lipschitz continuous curve $\mathbf{x}(\cdot) : [0, T] \to \mathbb{R}^n$, the corresponding classical *action* is

$$A_T[\mathbf{x}(\cdot)] := \int_0^T L(\dot{\mathbf{x}}, \mathbf{x})\, dt.$$

We call $\mathbf{x}(\cdot)$ a *critical point* of the action if it solves the *Euler-Lagrange* system of ODE $-\frac{d}{dt}\left(D_v L(\dot{\mathbf{x}}, \mathbf{x})\right) + D_x L(\dot{\mathbf{x}}, \mathbf{x}) = 0$. In this formula and below, $D_x := (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$, $D_v := (\frac{\partial}{\partial v_1}, \dots, \frac{\partial}{\partial v_n})$, etc.

Hamiltonian. We next define the *momentum* $\mathbf{p} := D_v L(\dot{\mathbf{x}}, \mathbf{x})$ and the *Hamiltonian* $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $H = H(p, x)$, by the formula

$$H(p,x) := \max_v \{p \cdot v - L(v,x)\}. \tag{1.3}$$

*See the Appendix for a translation between our notation and that more customary in physics and in dynamics.

Under reasonable assumptions, H is uniformly convex in the variable p.

The pair $\mathbf{x}, \mathbf{p}$ solves *Hamilton's equations*

$$(1.4) \qquad \begin{cases} \dot{\mathbf{x}} = D_p H(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}). \end{cases}$$

Generating function, canonical change of variables. To attempt to integrate these Hamiltonian ODE, we introduce a *generating function* $u : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $u = u(P, x)$. We propose a change of variables $(p, x) \to (P, X)$, implicitly defined by the formulas

$$(1.5) \qquad \begin{cases} p = D_x u(P, x) \\ X = D_P u(P, x). \end{cases}$$

Hamilton–Jacobi equations. Assume next u solves the stationary Hamilton–Jacobi type PDE

$$(1.6) \qquad H(D_x u(P, x), x) = \bar{H}(P) \quad \text{in } \mathbb{R}^n,$$

where at this point in the exposition the right hand side simply denotes some function of the variable P alone. Suppose as well that we can invert the relationships (1.5) to solve for P, X as smooth function of p, x. A calculation shows that we thereby transform (1.4) into the trivial dynamics

$$(1.7) \qquad \begin{cases} \dot{\mathbf{X}} = D\bar{H}(\mathbf{P}) \\ \dot{\mathbf{P}} = 0. \end{cases}$$

In the language of mechanics, P is an "action" and X an "angle" or "rotation" variable. See Goldstein [Gd] or Arnold–Kozlov–Neishtadt [A-K-N] for more.

Integrability and weak KAM theory. But we cannot carry out this classical procedure in general, since the PDE (1.6) does not usually admit a smooth solution and, even if it does, the canonical transformation $(p, x) \to (P, X)$ is not usually globally defined. Only very special Hamiltonians are integrable in this sense: see for instance Deift [D].

What remains is the possibility of coming up with some sort of weak interpretation of the classical program outlined above. Indeed Aubry[Au], Mather [Mt1-4], Fathi [F1-4], E [EW] and others have shown that certain solutions of (1.4), those arising from appropriate *minimizers* of the action, correspond to sorts of "integrable" structures within the full dynamics, generalizing the classical notion of invariant tori. Weak KAM theory (so named by A. Fathi) is the attempt to bring to bear global PDE techniques to continue this analysis, with particular emphasis upon problems with many degrees of freedom.

Let me emphasize also that, unlike conventional KAM theory, weak KAM theory is not perturbative. PDE and measure theory together provide us with solutions of the cell equation (C) and the transport equation (T), explained later, in the large. The fundamental technical issue is rather that these generalized solutions are not necessarily smooth, and so the classical calculations above are not obviously justified.

Some references. Good introductory lecture notes on Mather's variational principle in dynamics have been written by Contreras–Iturriaga [C-I] and by Fathi [F5]; and Forni–Mather [Fo-M] is another nice reference. See as well Mañé [Mn1-3], Fathi–Mather [Fa-M] and also Bernard [Be]. Different sorts of PDE methods are discussed in Jauslin–Kreiss–Moser [J-K-M] and E [EW].

Other recent developments of the PDE methods include these papers: Alvarez–Bardi [A-B], Barles–Souganidis [B-S], Concordel [C1,C2], [E-G2], Fathi–Maderna [F-Ma], Fathi–Siconolfi [F-S1,2], Gomes [G1-4], etc. See also Lions–Souganidis [L-S] for probabilistic methods and interpretations. One omission in this survey is any discussion of Fathi's PDE methods for "Peierls barriers" [F1-4].

1.2 An action minimization principle. Suppose that we have at hand a curve $\mathbf{x} : [0, \infty) \to \mathbb{R}^n$ that minimizes *for each time* $T > 0$ the action $A_T[\cdot]$, among Lipschitz curves $\mathbf{y}(\cdot)$ with $\mathbf{x}(0) = \mathbf{y}(0)$, $\mathbf{x}(T) = \mathbf{y}(T)$. We call such a curve an *absolute minimizer*, and are interested in absolute minimizers which satisfy for a given vector $V \in \mathbb{R}^n$ the asymptotic growth condition

$$\lim_{t\to\infty} \frac{\mathbf{x}(t)}{t} = V. \tag{1.8}$$

However in general absolute minimizing curves with these given asymptotics will not exist. This is shown by a famous example of Hedlund, discussed for instance in reference [EW].

A relaxed problem. As suggested by Mather, we can however "relax" the problem and look instead for a measure μ on the configuration space $\mathbb{R}^n \times \mathbb{T}^n$ that minimizes the *generalized action*

$$A[\mu] := \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(v, x)\, d\mu, \tag{1.9}$$

subject to the constraints that

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} v \cdot D\phi\, d\mu = 0 \quad \text{ for all } \phi \in C^1(\mathbb{T}^n), \tag{1.10}$$

$$\mu(\mathbb{R}^n \times \mathbb{T}^n) = 1, \quad \mu \geq 0, \tag{1.11}$$

and

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} v\, d\mu = V. \tag{1.12}$$

Condition (1.11) says of course that μ is a probability measure and (1.12) generalizes (1.8). Requirement (1.10), a "weak" flow invariance condition on μ, generalizes the classical idea that the action be computed along a curve. We will later reexamine this particular formulation of flow invariance, and will show that for a minimizer it implies an apparently stronger notion.

Mather has shown that in general there exist minimizing measures in this sense. We want to study their properties, in hopes of discovering some sort of "integrable structure" consistent with the foregoing classical formulation in terms of action-angle variables. For later reference, we define

$$\bar{L}(V) := \min_{\mu}\{A[\mu] \mid \text{conditions } (1.10) - (1.12) \text{ hold}\}. \tag{1.13}$$

The function $\bar{L} : \mathbb{R}^n \to \mathbb{R}$ so defined is the *effective Lagrangian*, and will reappear in another context later.

2. Linear programming insights.

Mather's variational interpretation in effect "linearizes" our problem: we are asked to minimize the linear functional (1.9), given the linear equality and inequality constraints (1.10) – (1.12). This is a linear programming problem in infinite dimensions, for which, as we show in this section, following [E-G3], a remarkable amount of useful information can be formally extracted from duality theory.

2.1 Review.

Finite dimensional linear programming. (Bertsimas–Tsitsiklis [B-T], Lax [L]) If $x \in \mathbb{R}^N$ is a vector, $x = (x_1, x_2, \dots, x_N)$, we write $x \geq 0$ to mean $x_i \geq 0$ for $i = 1, \dots, N$.

We are given vectors $c \in \mathbb{R}^N$, $b \in \mathbb{R}^M$ and an $M \times N$ matrix A. The *primal problem* is finding $\hat{x} \in \mathbb{R}^N$ to

$$\text{minimize } c \cdot x, \quad \text{subject to } Ax = b,\ x \geq 0; \tag{P}$$

and the corresponding *dual problem* is finding $\hat{y} \in \mathbb{R}^M$ to

$$\text{maximize } y \cdot b, \quad \text{subject to } A^T y \leq c. \tag{D}$$

Suppose that problem (P) has a solution $\hat{x}$ and (D) has a solution $\hat{y}$. A basic theorem of finite dimensional linear programming asserts that

$$c \cdot \hat{x} = \hat{y} \cdot b \tag{2.1}$$

or, equivalently,

$$\hat{x} \cdot (A^T \hat{y} - c) = 0. \tag{2.2}$$

This last equality expresses the *complementary slackness condition*, and implies for each $i = 1, \dots, N$, that either $\hat{x}_i = 0$ or $(A^T \hat{y} - c)_i = 0$, or both. Think of this as saying that if a constraint is inactive, the corresponding Lagrange multiplier is zero.

Infinite dimensional linear programming. (Anderson–Nash [A-N]) For our application to (1.9) – (1.12), we move to a more general setting. So suppose X, Y denote real vector spaces, and assume X is endowed with a partial ordering, inducing a partial ordering on its dual space X^*. The pairing between a space and its dual is denoted $\langle\, ,\, \rangle$.

Fix $c^* \in X^*$, $b \in Y$ and suppose $A : X \to Y$ is a bounded linear operator, with adjoint $A^* : Y^* \to X^*$. The *primal problem* is to find $\hat{x} \in X$ to

$$\text{(P)} \qquad \text{minimize } \langle c^*, x \rangle, \quad \text{subject to } Ax = b,\ x \geq 0;$$

and the *dual problem* is to find $\hat{y}^* \in Y^*$ to

$$\text{(D)} \qquad \text{maximize } \langle y^*, b \rangle, \quad \text{subject to } A^* y^* \leq c^*.$$

Suppose that $\hat{x}$ is optimal for (P) and $\hat{y}^*$ is optimal for (D). *If* it happens that

$$\langle \hat{y}^*, b \rangle = \langle c^*, \hat{x} \rangle, \tag{2.3}$$

then

$$\langle A^* \hat{y}^* - c^*, \hat{x} \rangle = 0; \tag{2.4}$$

and the identity (2.4) is another form of the complementary slackness condition.

2.2 Application to Mather's variational problem. Next we cast our variational problem into the foregoing framework and see in particular what complementary slackness implies. So let us take $X = \mathcal{M}(\mathbb{R}^n \times \mathbb{T}^n)$, the space of Radon measures on $\mathbb{R}^n \times \mathbb{T}^n$, and $Y = C^1(\mathbb{T}^n)^* \times \mathbb{R} \times \mathbb{R}^n$. Let $c^* = L \in X^*$, $b = (0^*, 1, V) \in Y$, where 0^* denotes the zero functional. If $\mu \in X$, we write

$$A\mu := \left(L_\mu, \mu(\mathbb{R}^n \times \mathbb{T}^n), \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} v \, d\mu \right). \tag{2.5}$$

In this expression L_μ denotes the linear functional defined for each $\phi \in C^1(\mathbb{T}^n)$ by the formula $L_\mu \phi := \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} v \cdot D\phi(x) \, d\mu$.

Primal problem. The primal problem (P) is therefore to find a measure μ to

$$\text{minimize } \langle c^*, \mu \rangle = \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} L \, d\mu, \tag{2.6}$$

subject to the requirements that

$$\mu \geq 0,\ A\mu = b. \tag{2.7}$$

Dual problem. Let $y^* = (v(\cdot), w_0, w) \in Y^* = C^1(\mathbb{T}^n)^{**} \times \mathbb{R} \times \mathbb{R}^n$. We will vastly simplify this heuristic discussion, by *just assuming hereafter that in fact $v(\cdot)$ is a C^1 function.* (There is a notational pitfall here, since v also denotes the velocity variable; and I will therefore carefully distinguish below between the variable $v \in \mathbb{R}^n$ and the function $v(\cdot) \in C^1(\mathbb{T}^n)$. Accepting this for the moment resolves many notational problems later.)

We compute A^*y^* by calculating for $x = \mu$ that

$$\langle A^*y^*, x\rangle = \langle y^*, Ax\rangle = \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} v \cdot Dv(x) + w_0 + w \cdot v \, d\mu,$$

and deduce

$$A^*y^* = w_0 + v \cdot (w + Dv(x)). \tag{2.8}$$

So the dual problem is to find $y^* = (v(\cdot), w_0, w)$, to

$$\text{maximize } \ \langle y^*, b\rangle = w_0 + V \cdot w, \tag{2.9}$$

subject to the pointwise constraints that

$$w_0 + v \cdot (w + Dv(x)) \le L(v, x) \quad \text{in } \mathbb{R}^n \times \mathbb{T}^n. \tag{2.10}$$

2.3 Interpretations of complementary slackness. We now assume that $\hat{x} = \mu$ and $\hat{y}^* = (v(\cdot), w_0, w)$ are optimal. We continue simply to suppose that $v(\cdot) \in C^1$, and also for heuristic purposes hypothesize that the complementary slackness condition is valid. This says

$$w_0 + v \cdot (w + Dv(x)) = L(v, x) \quad \text{on } \operatorname{spt}(\mu). \tag{2.11}$$

I write "spt" for support.

Geometry of the support of $\boldsymbol{\mu}$. For each fixed $x \in \mathbb{T}^n$, define the function

$$\psi(v) := L(v, x) - v \cdot (w + Dv(x)) - w_0 \quad (v \in \mathbb{R}^n).$$

Then

$$\psi \ge 0 \ \text{ on } \mathbb{R}^n \ \text{ and } \ \psi = 0 \ \text{ on } \operatorname{spt}(\mu). \tag{2.12}$$

Since L and therefore ψ are uniformly convex in the variable v, it follows for each $x \in \mathbb{T}^n$ that the support of μ in $\mathbb{R}^n \times \{x\}$ consists of at most one point $v \in \mathbb{R}^n$. Since the gradient of ψ vanishes at its minimum, we in fact have $D_vL(v, x) = w + Dv(x)$. Consequently, *the support of the measure μ lies on the n-dimensional graph $v = (D_vL(\cdot, x))^{-1}(w + Dv(x))$ over $\mathbb{T}^n$*. This is a sort of regularity assertion for the structure of the minimizing measure μ.

Solutions of a generalized eikonal PDE. Owing to (2.10), we have

$$H(w + Dv(x), x) = \max_{v \in \mathbb{R}^n} \{v \cdot (w + Dv(x)) - L(v, x)\} \le -w_0 \quad \text{on } \mathbb{T}^n, \tag{2.13}$$

with equality on the support of $\sigma := \operatorname{proj}_x \mu$, the projection of μ onto the x-variables. We change notation, now to write

$$P := w, \quad \bar{H}(P) := -w_0. \tag{2.14}$$

Now writing $Dv = Dv(x)$, we have

$$H(P + Dv, x) \leq \bar{H}(P) \quad \text{on } \mathbb{T}^n, \tag{2.15}$$

with

$$H(P + Dv, x) = \bar{H}(P) \quad \text{on } \operatorname{spt}(\sigma). \tag{2.16}$$

The support of our measure consequently lies within the graph $v = D_pH(P + Dv(x), x) = D_pH(Du(x), x)$, for

$$u := P \cdot x + v. \tag{2.17}$$

Weak flow invariance implies flow invariance. Next we show that the weak flow invariance condition (1.10) implies actual flow invariance for our minimizer μ. To do so, it is most convenient to switch to the Hamiltonian formulation. We consequently use the mapping $p = D_vL(v, x)$ to push forward μ from the configuration space with variables (v, x) to a measure ν on the phase space with variables (p, x). Since $v = D_pH(p, x)$, (1.10) says for $\phi \in C^1(\mathbb{T}^n)$, $\phi = \phi(x)$, that

$$0 = \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} v \cdot D\phi \, d\mu = \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_pH(Du, x) \cdot D\phi \, d\nu.$$

Now take $\Phi \in C^1(\mathbb{R}^n \times \mathbb{T}^n)$, $\Phi = \Phi(p, x)$, and write $\phi(x) := \Phi(Du(x), x)$. Then $D\phi = D_x\Phi + D_p\Phi D^2u$. Furthermore (2.15), (2.16) imply that $D_xH + D_pHD^2u = 0$ on the support of σ. Therefore, letting $\{\cdot\}$ denotes the Poisson bracket, we deduce

$$\begin{aligned}
\int_{\mathbb{R}^n} \int_{\mathbb{T}^n} \{H, \Phi\} \, d\nu &= \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_pH(Du(x), x) \cdot D_x\Phi - D_xH(Du(x), x) \cdot D_p\Phi \, d\nu \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_pH(Du(x), x) \cdot D_x\Phi + D_pH(Du(x), x)D^2u \cdot D_p\Phi \, d\nu \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_pH(Du(x), x) \cdot D\phi \, d\nu = 0.
\end{aligned}$$

This identity says that ν is invariant under the full Hamiltonian flow (1.4). This is a variant of an observation due originally to Mañé.

What is linear programming telling us? The foregoing formalism suggests strongly that we focus attention on the *generalized eikonal equation* (2.16), the right hand side of which is an unknown function of the variable P alone. Recalling (2.17), we note that this is formally the same as (1.6).

For future reference, we restate the eikonal equation:

$$H(Du, x) = \bar{H}(P) \quad \text{on } \operatorname{spt}(\sigma), \tag{2.18}$$

the right hand side of which will turn out to be the *effective Hamiltonian* (about which more later) evaluated at $P \in \mathbb{R}^n$. Furthermore, we can check that $\sigma = \operatorname{proj}_x \mu = \operatorname{proj}_x \nu$ solves the *transport* (or *continuity*) *equation*

$$\operatorname{div}(\sigma D_p H(Du, x)) = 0. \tag{2.19}$$

The point is that the formalism of linear programming has identitified these PDE as being somehow relevant. It suggests also these interpretations:

(i) The vector P is the Lagrange multiplier for the constraint (1.12) that the measure μ have the rotation, or average velocity, vector V.

(ii) The number $\bar{H}(P)$ is the Lagrange multiplier for the constraint (1.11) that μ be a probability measure.

(iii) The function u, solving the generalized eikonal equation, is the Lagrange multiplier for the constraint (1.10) that μ be weakly flow invariant.

We emphasize that all the calculations and conclusions of this section are purely formal: our goal now is extracting some rigorous theory. But it is interesting, and I think surprising, that the "soft" principles of linear programming predict such detailed structure.

3. The effective Hamiltonian.

3.1 How to construct $\bar{\mathbf{H}}$. We can rigorously build appropriate *weak* solutions of (2.18), as shown in the unpublished, but classic, paper Lions–Papanicolaou–Varadhan [L-P-V] . Consider for fixed $P \in \mathbb{R}^n$ the *cell problem*

$$\begin{cases} H(P + Dv, x) = \lambda & \text{in } \mathbb{R}^n, \\ x \mapsto v \ \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases}$$

As proved in [L-P-V] (and recounted in [E1]), there exists a unique real number λ for which there exists a viscosity solution $v = v(x)$, in the sense of Crandall–Lions. We may then *define* $\bar{H}(P) := \lambda$ and set

$$u := P \cdot x + v, \tag{3.1}$$

to recast the foregoing as

$$H(Du, x) = \bar{H}(P) \quad \text{in } \mathbb{R}^n. \tag{C}$$

The label "C" stands for either "cell problem" or "corrector problem". This PDE is of course a rigorously derived form of (2.18).

We understand v to solve (C) in the sense of viscosity solutions, but actually will mostly need only that v is differentiable a.e. and that v solves the PDE at any point of differentiability. It turns out that the function u is semiconcave, meaning that $u_{\xi\xi} \leq \alpha$ in the sense of distributions, for all unit vectors ξ and some constant α. We will sometimes think of u, v as depending only upon x, with P fixed, and will sometimes instead imagine u, v as depending upon both x and P.

Let us call the function $\bar{H} : \mathbb{R}^n \to \mathbb{R}$ so defined the *effective Hamiltonian*. Given $\bar{H}$ as above, we define also the *effective Lagrangian*

$$\bar{L}(V) := \sup_P (P \cdot V - \bar{H}(P)) \tag{3.2}$$

for $V \in \mathbb{R}^n$. Later we will show that this agrees with our previous definition (1.13). The mappings $\bar{H}$ and $\bar{L}$ are convex, real-valued and superlinear.

Numerical computations. As we will see in following sections, the idea is that the effective Hamiltonian $\bar{H}$ "encodes" information about certain dynamics: it is consequently important to understand the structure of this function of P. It has however proved difficult to obtain explicit formulas in dimensions $n \geq 2$, and hence numerical calculations seem useful.

Gomes and Oberman [G-O] and Qian [Q] have numerically computed $\bar{H}$ for $n = 2$ in several interesting cases, including $H(p, x) = \frac{1}{2}|p|^2 + W(x)$. See also Bourlioux–Khouider [B-K], Oberman [O], etc.

3.2 Minimizing measures. ([E-G1]) Having now at hand a weak solution of the eikonal equation, we want to use it to study the minimizing measure μ. To do so, it is most convenient to switch as in §2 to the Hamiltonian formulation. We therefore assume we have a compactly supported Radon probability measure ν on phase space $\mathbb{R}^n \times \mathbb{T}^n$, for which

$$V = \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_p H(p, x) \, d\nu \tag{3.3}$$

and

$$\int_{\mathbb{R}^n} \int_{\mathbb{T}^n} \{H, \Phi\} \, d\nu = 0 \tag{3.4}$$

for each C^1 function Φ that is $\mathbb{T}^n$-periodic, where $\{\cdot\}$ again is the Poisson bracket. We also suppose

$$\bar{L}(V) = \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} L(D_p H(p, x), x) \, d\nu. \tag{3.5}$$

We assume afterwards that the Hamiltonian $H = H(p, x)$ is $\mathbb{T}^n$-periodic in x and uniformly convex in p, meaning that $D_p^2 H \geq \gamma I$ as symmetric matrices for some constant $\gamma > 0$. Note that $H(p, x) + L(v, x) \geq p \cdot v$, with equality if and only if $p = D_v L(v, x), v = D_p H(p, x)$. The projection of ν onto the x–variables is denoted σ.

Take now any $P \in \partial \bar{L}(V)$ and let $u = P \cdot x + v$ be any viscosity solution of the corresponding cell problem (C).

Theorem 3.1. (i) *The function u is differentiable in the variable x σ-a.e., and σ-a.e. point is a Lebesgue point for $D_x u$.*

(ii) *We have*

$$p = D_x u(P, x) \quad \nu\text{-a.e.} \tag{3.6}$$

(iii) *Furthermore,*

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} H(p,x)\,d\nu = \int_{\mathbb{T}^n} H(D_x u, x)\,d\sigma = \bar{H}(P); \tag{3.7}$$

and if $\bar{H}$ is differentiable at P,

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_p H(p,x)\,d\nu = \int_{\mathbb{T}^n} D_p H(D_x u, x)\,d\sigma = D\bar{H}(P). \tag{3.8}$$

Thus ν is supported on the graph $p = D_x u(P,x) = P + D_x v(P,x)$, which is single-valued σ-a.e. Also, the corrector PDE (C) holds pointwise, σ-a.e. Compare assertion (ii) with the classical canonical change of variables (1.5). Also,

$$\operatorname{div}(\sigma D_p H(D_x u, x)) = 0 \quad \text{in } \mathbb{T}^n, \tag{T}$$

the label "T" standing for "transport equation"; and this is identical with the PDE (2.19) formally derived earlier.

Idea of proof. We do not display the dependence of u on the variable P, and also write Du for $D_x u$.

Take η_ε to be a smooth, nonnegative, radial convolution kernel, supported in the ball $B(0,\varepsilon)$, and set $u^\varepsilon := \eta_\varepsilon * u$. Using the uniform convexity of H in p and Jensen's inequality, we find that

$$\beta_\varepsilon(x) + H(Du^\varepsilon(x), x) \le \bar{H}(P) + C\varepsilon \tag{3.9}$$

for each $x \in \mathbb{T}^n$, where

$$\beta_\varepsilon(x) := \frac{\gamma}{2}\int_{\mathbb{R}^n} \eta_\varepsilon(x-y)|Du(y) - Du^\varepsilon(x)|^2\,dy \tag{3.10}$$

for some constant $\gamma > 0$. Recalling again the strict convexity of H, we have

$$\begin{aligned}&\frac{\gamma}{2}\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} |Du^\varepsilon(x) - p|^2\,d\nu \\ &\qquad \le \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} H(Du^\varepsilon(x), x) - H(p,x) - D_p H(p,x)\cdot(Du^\varepsilon(x) - p)\,d\nu.\end{aligned} \tag{3.11}$$

Now $Du^\varepsilon = P + Dv^\varepsilon$, where $v^\varepsilon = \eta_\varepsilon * v$ is periodic. Consequently $\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_pH\cdot Dv^\varepsilon\, d\nu = 0$, according to (3.4). Then

$$\begin{aligned}\frac{\gamma}{2}\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} |Du^\varepsilon - p|^2\, d\nu + \int_{\mathbb{T}^n}\beta_\varepsilon\, d\sigma \\ \le \bar{H}(P) - \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} H + D_pH\cdot(P-p)\, d\nu + C\varepsilon.\end{aligned} \tag{3.12}$$

Next, $P \in \partial\bar{L}(V)$ implies $\bar{L}(V) + \bar{H}(P) = P\cdot V$. Furthermore $L(D_pH(p,x),x) + H(p,x) = D_pH(p,x)\cdot p$. Recalling that $V = \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_pH\, d\nu$ and substituting into (3.12), we find

$$\begin{aligned}\frac{\gamma}{2}\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} |Du^\varepsilon - p|^2\, d\nu + \int_{\mathbb{T}^n}\beta_\varepsilon\, d\sigma \\ \le -\bar{L}(V) + \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(D_pH,x)\, d\nu + C\varepsilon = C\varepsilon.\end{aligned} \tag{3.13}$$

Now send $\varepsilon \to 0$. Passing as necessary to a subsequence, we deduce first that $\beta_\varepsilon \to 0$ σ-a.e. Thus σ-a.e. point x is a point of approximate continuity of Du. Since u is semiconcave as a function of x, it follows that u is differentiable in x, σ-a.e. Hence $Du^\varepsilon \to Du$ pointwise, σ-a.e., and so (3.13) also forces $p = Du(x) = P + Dv(x)$ ν-a.e. □

3.3 A minimax formula. The following characterization of the effective Hamiltonian will be useful later:

Theorem 3.2. *We have*

$$\bar{H}(P) = \inf_{v\in C^1(\mathbb{T}^n)} \max_{x\in\mathbb{T}^n} H(P+Dv,x). \tag{3.14}$$

Proof. Let u be a viscosity solution of $H(Du,x) = \bar{H}(P)$ and let σ be a corresponding measure; so that $\operatorname{div}(\sigma D_pH(Du,x)) = 0$. As before, we think of u as depending only on x. Take $\hat{u} = P\cdot x + \hat{v}$, where $\hat{v}$ any $\mathbb{T}^n$-periodic, C^1 function. Then convexity implies $H(Du,x) + D_pH(Du,x)\cdot(D\hat{u} - Du) \le H(D\hat{u},x)$. We integrate with respect to σ:

$$\bar{H}(P) = \int_{\mathbb{T}^n} H(Du,x)\, d\sigma \le \int_{\mathbb{T}^n} H(D\hat{u},x)\, d\sigma \le \max_{x\in\mathbb{T}^n} H(D\hat{u},x).$$

Thus $\bar{H}(P) \le \inf_{\hat{v}\in C^1}\max_{x\in\mathbb{T}^n} H(P+D\hat{v},x)$.

On the other hand, the smoothed function $u^\varepsilon := \eta_\varepsilon * u = P\cdot x + v^\varepsilon$ satisfies as above $H(Du^\varepsilon,x) \le \bar{H}(P) + O(\varepsilon)$ uniformly. And so $\inf_{\hat{v}\in C^1}\max_{x\in\mathbb{T}^n} H(P+D\hat{v},x) \le \bar{H}(P)$. □

This quick proof is due to A. Fathi. Several authors, among them Mañé, Contreras–Iturriaga–Paternain and Gomes, have independently derived this identity. Fathi and Siconolfi [F-S] have recently shown that there exists a C^1 subsolution of (C).

3.4 Homogenization. The effective Hamiltonian first arose in the PDE literature in consideration of periodic homogenization problems for Hamilton–Jacobi equations, as introduced by Lions–Papanicolaou–Varadhan [L-P-V]. (See also [E-1].) These authors look at the initial value problem:

$$
\begin{cases} u^\varepsilon_t + H\left(Du^\varepsilon, \frac{x}{\varepsilon}\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u^\varepsilon = g & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases} \tag{3.15}
$$

under the primary assumption that the mapping $x \mapsto H(p,x)$ is $\mathbb{T}^n$-periodic. Consequently as $\varepsilon \to 0$ the nonlinearity in (3.15) is rapidly oscillating; and the problem is to understand the limiting behavior of the solutions u^ε. Lions, Papanicolaou and Varadhan show that $u^\varepsilon \to u$, the limit function u solving the intial value problem

$$
\begin{cases} u_t + \bar{H}(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}
$$

Majda and Souganidis [M-S] discuss an interesting application of these ideas to turbulent premixed flames: see also Bourlioux–Khouider [B-K] for numerical studies.

Convexity of H. It is worth remarking that Lions *et al.* do not need to assume that $p \mapsto H(p,x)$ be convex, and indeed their solution of the cell problem (C) requires only the coercivity condition that $H(p,x) \to \infty$ as $p \to \infty$. It is a major open problem to interpret what $\bar{H}$ means for dynamics, should H be nonconvex in the momenta.

4. Partial regularity theory.

We devote this section to showing as in [E-G1] that our solution u of the cell problem is "smoother" on the support of σ than it may be elsewhere in $\mathbb{T}^n$.

4.1 Derivative estimates in the variable x. We provide first some purely formal L^2 and L^∞ estimates for $D^2_x u$ on the support of σ.

***L^2*-inequalities**. We assume that u is smooth, differentiate the cell PDE (C) twice with respect to x_i, and add:

$$
\begin{aligned} H_{p_k p_l}(D_x u, x) u_{x_k x_i} u_{x_l x_i} &+ H_{p_k}(D_x u, x) u_{x_k x_i x_i} \\ &+ 2H_{p_k x_i}(D_x u, x) u_{x_k x_i} + H_{x_i x_i}(D_x u, x) = 0. \end{aligned}
$$

By uniform convexity, the first term on the left is greater than or equal to $\gamma |D^2_x u|^2$. Thus

$$
\gamma \int_{\mathbb{T}^n} |D^2_x u|^2 \, d\sigma + \int_{\mathbb{T}^n} D_p H \cdot D_x(\Delta_x u) \, d\sigma \leq C + C \int_{\mathbb{T}^n} |D^2_x u| \, d\sigma.
$$

Since $\Delta_x u = \Delta_x v$ is periodic, the second term on the left equals zero, according to (T). We consequently conclude

$$
\int_{\mathbb{T}^n} |D^2_x u|^2 \, d\sigma \leq C, \tag{4.1}
$$

for some constant C depending only on H and P. □

***L^∞*-inequalities**. We can similarly differentiate the cell PDE twice in any unit direction ξ, to find

$$H_{p_k p_l}(D_x u, x) u_{x_k \xi} u_{x_l \xi} + H_{p_k}(D_x u, x) u_{x_k \xi\xi} \\ + 2H_{p_k \xi}(D_x u, x) u_{x_k \xi} + H_{\xi\xi}(D_x u, x) = 0,$$

for $u_{\xi\xi} := \sum_{i,j=1}^n u_{x_i x_j} \xi_i \xi_j$. Take a nondecreasing, function $\Phi : \mathbb{R} \to \mathbb{R}$, and write $\phi := \Phi' \geq 0$. Multiply the above identity by $\phi(u_{\xi\xi})$, and integrate with respect to σ. After some simplifications, we find

$$\frac{\gamma}{2} \int_{\mathbb{T}^n} |D_x u_\xi|^2 \phi(u_{\xi\xi})\, d\sigma + \int_{\mathbb{T}^n} D_p H \cdot D_x(\Phi(u_{\xi\xi}))\, d\sigma \leq C \int_{\mathbb{T}^n} \phi(u_{\xi\xi})\, d\sigma.$$

Since $u_{\xi\xi} = v_{\xi\xi}$ is periodic, the second term on the left is zero. We select $\phi(z) = 1$ if $z \leq -\beta$ and $\phi(z) = 0$ if $z > -\beta$, for a constant $\beta > 0$. Since $|D_x u_\xi|^2 \geq u_{\xi\xi}^2$, we conclude that $\sigma(\{u_{\xi\xi} \leq -\beta\}) = 0$ if μ is large enough. Because semiconcavity provides the opposite estimate $u_{\xi\xi} \leq \alpha$, we thereby derive the formal bound

$$|u_{\xi\xi}| \leq C \quad \sigma\text{-a.e.}, \tag{4.2}$$

the constant C depending only upon known quantities. □

We next state a rigorous analog of estimates (4.1), (4.2), with difference quotients replacing some of the derivatives.

Theorem 4.1. (i) *There exists a constant C, depending only on H and P, such that*

$$\int_{\mathbb{T}^n} |D_x u(P, x+h) - D_x u(P, x)|^2\, d\sigma \leq C|h|^2 \tag{4.3}$$

for $h \in \mathbb{R}^n$.

(ii) *In addition, there exists a constant C such that*

$$|u(P, x+h) - 2u(P, x) + u(P, x-h)| \leq C|h|^2 \tag{4.4}$$

for all $h \in \mathbb{R}^n$ and each point $x \in \operatorname{spt}(\sigma)$.

If $D_x u(P, x+h)$ is multivalued, we interpret (4.3) to mean

$$\int_{\mathbb{T}^n} |\boldsymbol{\xi} - D_x u|^2\, d\sigma \leq C|h|^2 \tag{4.5}$$

for some σ-measurable selection $\boldsymbol{\xi} \in D_x u(P, \cdot + h)$.

Application: Lipschitz estimates for the support of $\boldsymbol{\nu}$. We next note that $\operatorname{spt}(\nu)$ lies on a Lipschitz continuous graph.

Theorem 4.2. *There exists a constant C such that*

$$|u(P,y) - u(P,x) - D_x u(P,x) \cdot (y-x)| \leq C|x-y|^2 \tag{4.6}$$

for all $y \in \mathbb{T}^n$ and σ-a.e. point $x \in \mathbb{T}^n$. Furthermore,

$$|D_x u(P,y) - D_x u(P,x)| \leq C|x-y| \tag{4.7}$$

for all $y \in \mathbb{T}^n$ and for σ-a.e. point $x \in \mathbb{T}^n$.

In fact, u is differentiable at each point $x \in \operatorname{spt}(\sigma)$, and estimates (4.6), (4.7) hold for all $y \in \mathbb{T}^n, x \in \operatorname{spt}(\sigma)$. When $D_x u(P,y)$ is multivalued, (4.7) asserts $|\xi - D_x u(P,x)| \leq C|x-y|$ for all $\xi \in D_x u(P,y)$. In particular, for multivalued $D_x u(P,y)$ we have the estimate $\operatorname{diam}(D_x u(P,y)) \leq C \operatorname{dist}(y, \operatorname{spt}(\sigma))$, providing a quantitative justification to the informal assertion that "the support of σ misses the shocks in Du".

As an application of these bounds, we observe next that the set $\mathcal{M} = \operatorname{spt}(\nu)$ lies on an n-dimensional Lipschitz continuous graph. This important theorem (in position-velocity variables) is due to Mather.

Theorem 4.3. ([Mt2]) *There exists a constant C for which*

$$|D_x u(P,x_1) - D_x u(P,x_2)| \leq C|x_1 - x_2| \tag{4.8}$$

for σ-a.e. pair of points x_1, x_2.

4.2 Derivative estimates in the variable P. We turn next to some bounds involving variations in P, and as before begin with a heuristic calculation. So for the moment suppose u and $\bar{H}$ are smooth, differentiate the cell PDE (C) twice with respect to P_i, and sum:

$$H_{p_k p_l}(D_x u, x) u_{x_k P_i} u_{x_l P_i} + H_{p_k}(D_x u, x) u_{x_k P_i P_i} = \bar{H}_{P_i P_i}(P).$$

The first term is greater than or equal to $\gamma |D^2_{xP} u|^2$. Consequently

$$\gamma \int_{\mathbb{T}^n} |D^2_{xP} u|^2 \, d\sigma + \int_{\mathbb{T}^n} D_p H \cdot D_x(\Delta_{\scriptscriptstyle P} u) \, d\sigma \leq \Delta \bar{H}(P),$$

where $\Delta \bar{H}$ is the Laplacian of $\bar{H}$ in P. Since $\Delta_{\scriptscriptstyle P} u = \Delta_{\scriptscriptstyle P} v$ is periodic, (T) implies the second term on the left equals zero. Therefore

$$\int_{\mathbb{T}^n} |D^2_{xP} u|^2 \, d\sigma \leq C \Delta \bar{H}(P) \leq C, \tag{4.9}$$

if $D^2 \bar{H}(P)$ exists and is bounded. (I do not know how even formally to derive a sup-norm bound for $|D^2_{xP} u|$ on the support of σ.)

We next state a rigorous version of this calculation, which replaces derivatives by difference quotients.

Theorem 4.4. *There exists a positive constant C such that*

$$\int_{\mathbb{T}^n} |D_x u(\tilde{P}, x) - D_x u(P, x)|^2 \, d\sigma \le C(\bar{H}(\tilde{P}) - \bar{H}(P) - V \cdot (\tilde{P} - P)) \tag{4.10}$$

for all $\tilde{P} \in \mathbb{R}^n$.

If $D_x u(\tilde{P}, x)$ is multivalued, we interpret (4.10) to mean

$$\int_{\mathbb{T}^n} |\tilde{\boldsymbol{\xi}} - D_x u(P, x)|^2 \, d\sigma \le C(\bar{H}(\tilde{P}) - \bar{H}(P) - V \cdot (\tilde{P} - P))$$

for some σ-measurable selection $\tilde{\boldsymbol{\xi}} \in D_x u(\tilde{P}, \cdot)$.

Application: strict convexity of $\bar{H}$ in certain directions. The next estimate allows us to deduce certain strict convexity properties of $\bar{H}$.

Theorem 4.5. *There exists a positive constant C such that if $\bar{H}$ is twice differentiable at P, then*

$$|D\bar{H}(P) \cdot R| \le C(R \cdot D^2 \bar{H}(P) R)^{1/2} \tag{4.11}$$

for each $R \in \mathbb{R}^n$.

So if $D^2 \bar{H}(P)$ exists, the effective Hamiltonian is strictly convex in any direction R which is not tangent to its level set $\{\bar{H} = \bar{H}(P)\}$.

Idea of proof. We provide only the relevant formal calculations here. Differentiating the cell problem with respect to P gives $D_p H D_{xP} u = D\bar{H}$. Therefore

$$|D\bar{H}(P) \cdot R| \le C \int_{\mathbb{T}^n} |D_{xP} u \cdot R| \, d\sigma \le C \left(\int_{\mathbb{T}^n} |D_{xP} u \cdot R|^2 \, d\sigma \right)^{\frac{1}{2}} \le C(R \cdot D^2 \bar{H}(P) R)^{1/2},$$

the last inequality following from calculations as above. See [E-G1] for details. □

5. A new variational principle.

5.1 Calculus of variations in the sup–norm. The minimax formula (3.14) suggests that we can compute $\bar{H}(P)$ by trying to minimize the sup–norm of $H(P + Dv, x)$ over $\mathbb{T}^n$. This viewpoint is strongly reminiscent of "Aronsson's variational principle" in the calculus of variations: see Barron [B] for more about this. The idea is that we should not just try to minimize the sup–norm of $H(P + Dv, x)$ over $\mathbb{T}^n$, but should rather look for a function v which minimizes

$$||H(P + Dv, x)||_{L^\infty(U)},$$

relative to its boundary values, for each open subdomain $U \subset \mathbb{T}^n$. We then call v an *absolute minimizer.*

In this section we take inspiration from these ideas, to construct a new approximation using an exponential expression, as in [E3]. An advantage is that this technique will simultaneously build a (sub)solution u of the eikonal equation (C) and a measure solution σ of the transport equation (T). So given a positive integer k, we look for $v^k \in C^1(\mathbb{T}^n)$ to minimizing the functional

$$I_k[v] := \int_{\mathbb{T}^n} e^{kH(P+Dv^k,x)}\,dx. \tag{5.1}$$

Under usual assumptions on H, there exists a minimizer $v^k \in C^\infty(\mathbb{T}^n)$, which is unique once we require $\int_{\mathbb{T}^n} v^k\,dx = 0$.

PDE interpretations. The Euler–Lagrange equation for our minimizer of $I_k[\cdot]$ is

$$\operatorname{div}(e^{kH(Du^k,x)}D_pH(Du^k,x)) = 0, \tag{5.2}$$

where we have as usual set $u^k := P\cdot x + v^k$. Define

$$\sigma^k := \frac{e^{kH(Du^k,x)}}{\int_{\mathbb{T}^n} e^{kH(Du^k,x)}dx} = e^{k(H(Du^k,x)-\bar{H}^k(P))}, \tag{5.3}$$

for

$$\bar{H}^k(P) := \frac{1}{k}\log\left(\int_{\mathbb{T}^n} e^{kH(Du^k,x)}dx\right). \tag{5.4}$$

Then $\sigma^k \geq 0$ and $\int_{\mathbb{T}^n} d\sigma^k = 1$, where $d\sigma^k := \sigma^k\,dx$. Observe that the Euler–Lagrange equation (5.2) now reads

$$\operatorname{div}(\sigma^k D_pH(Du^k,x)) = 0. \tag{5.5}$$

This is a form of our transport equation (T).

5.2 Convergence, approximating the effective Hamiltonian. Passing if necessary to a subsequence, we have $u^k \to u = P\cdot x + v$ uniformly, and $Du^k \rightharpoonup Du = P + Dv$ weakly in $L^v(\mathbb{T}^n;\mathbb{R}^n)$, where $v \in W^{1,\infty}(\mathbb{T}^n)$. In addition, we may also suppose that $\sigma^k \rightharpoonup \sigma$ weakly as measures, where σ is a Radon probability measure on $\mathbb{T}^n$.

Theorem 5.1. (i) *We have*

$$\bar{H}(P) = \lim_{k\to\infty} \bar{H}^k(P) = \lim_{k\to\infty}\frac{1}{k}\log\left(\int_{\mathbb{T}^n} e^{kH(Du^k,x)}dx\right). \tag{5.6}$$

(ii) *The function u is a viscosity solution of* Aronsson's equation

$$-H_{p_i}(Du,x)H_{p_j}(Du,x)u_{x_ix_j} = H_{x_i}(Du,x)H_{p_i}(Du,x). \tag{5.7}$$

(iii) *Furthermore,*

$$H(Du, x) \leq \bar{H}(P) \qquad \textit{a.e. in } \mathbb{T}^n. \tag{5.8}$$

The mysterious PDE (5.7) is in effect the Euler-Lagrange equation for our limit "variational problem in the sup-norm". Several authors have proposed so naming this equation to honor G. Aronsson's pioneering contributions.

Idea of proof. We will only prove (i). Take $u = P \cdot x + v$ solving (C). Then

$$\int_{\mathbb{T}^n} e^{kH(P+Dv^k,x)} dx \leq \int_{\mathbb{T}^n} e^{kH(P+Dv,x)} dx = e^{k\bar{H}(P)},$$

and consequently

$$\limsup_{k\to\infty} \frac{1}{k} \log \left(\int_{\mathbb{T}^n} e^{kH(Du^k,x)} dx \right) \leq \bar{H}(P).$$

Suppose next that

$$\liminf_{k\to\infty} \frac{1}{k} \log \left(\int_{\mathbb{T}^n} e^{kH(Du^k,x)} dx \right) < \bar{H}(P) - \varepsilon$$

for some $\varepsilon > 0$. Let $L_{\varepsilon,k} := \left\{ x \in \mathbb{T}^n \mid H(Du^k, x) > \bar{H}(P) - \frac{\varepsilon}{2} \right\}$. Then for some sequence $k_j \to \infty$, we have $\frac{1}{k_j} \log |L_{\varepsilon,k_j}| + \bar{H}(P) - \frac{\varepsilon}{2} \leq \bar{H}(P) - \varepsilon$. This implies $|L_{\varepsilon,k_j}| \leq e^{-k_j \frac{\varepsilon}{2}}$. Hence for each measurable set $A \subset \mathbb{T}^n$ with $|A| > 0$,

$$⨍_A H(Du, x)\, dx \leq \liminf_{k_j\to\infty} ⨍_A H(Du^{k_j}, x)\, dx \leq \bar{H}(P) - \frac{\varepsilon}{2},$$

the slash through the integral denoting an average, and so $H(Du, x) \leq \bar{H}(P) - \frac{\varepsilon}{2}$ a.e.

We now set $u^\delta := \eta_\delta * u$, where η_δ denotes a standard mollifier. Then $H(Du^\delta, x) \leq \bar{H}(P) - \frac{\varepsilon}{4}$ in $\mathbb{T}^n$ for sufficiently small $\delta > 0$. This however contradicts the minimax formula (3.14), as $u^\delta = P \cdot x + v^\delta$ for some smooth, periodic function v^δ. $\square$

5.3 Minimizing measures again. To understand more about the structure of the measure σ, it is convenient to lift into $\mathbb{R}^n \times \mathbb{T}^n$, as follows. Define

$$\mu^k := \delta_{\{v = D_pH(Du^k,x)\}} \sigma^k; \tag{5.9}$$

that is, $\int_{\mathbb{R}^n} \int_{\mathbb{T}^n} \Phi(v, x)\, d\mu^k = \int_{\mathbb{T}^n} \Phi(D_pH(Du^k, x), x)\, d\sigma^k$ for each C^1 function Φ.

It is not hard to check that the family of probability measures $\{\mu^k\}_{k=1}^\infty$ on $\mathbb{R}^n \times \mathbb{T}^n$ is tight. Passing to a subsequence if necessary, we may suppose that $\mu^k \rightharpoonup \mu$ weakly as measures, for μ a probability measure on $\mathbb{R}^n \times \mathbb{T}^n$. Note that $\sigma = \mathrm{proj}_x\, \mu$, the projection of μ into $\mathbb{T}^n$. Define also the vector

$$V := \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} v\, d\mu. \tag{5.10}$$

Theorem 5.2. (i) *The measure* μ *is weakly flow invariant; that is,*

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} v\cdot D\phi\, d\mu = 0 \tag{5.11}$$

for all $\phi \in C^1(\mathbb{T}^n)$.

(ii) *The limit*

$$\lim_{k\to\infty}\int_{\mathbb{T}^n} H(Du^k, x)\, d\sigma^k = \bar{H}(P) \tag{5.12}$$

holds.

(iii) *Furthermore,*

$$A[\mu] = \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(v,x)\, d\mu = \bar{L}(V), \quad V \in \partial\bar{H}(P). \tag{5.13}$$

(iv) *The function* u *is differentiable* σ*-almost everywhere and*

$$v = D_pH(Du,x) \qquad \mu\text{-almost everywhere.} \tag{5.14}$$

In particular, $V := \int_{\mathbb{T}^n} D_pH(Du,x)\, d\sigma$.

(v) *We have*

$$H(Du,x) = \bar{H}(P) \qquad \sigma\text{-almost everywhere} \tag{5.15}$$

and

$$\operatorname{div}(\sigma D_pH) = 0 \quad \text{in } \mathbb{T}^n. \tag{5.16}$$

Note we are only asserting that the eikonal PDE (5.15) holds on the support of σ. This is the primary difference between our approach and that of §3-4. Equations (5.15) and (5.16) are forms of our basic PDE (C) and (T).

Remark: the two definitions of $\bar{\mathbf{L}}$. We can now show that the effective Lagrangian $\bar{L}$, defined in (3.12) as the convex dual of $\bar{H}$, agrees with the earlier definition (1.13) (as the value of Mather's variational problem).

For this, let $\tilde{\mu}$ be any measure satisfying (1.10) − (1.12) and suppose temporarily that $u = P\cdot x + v$ built above is smooth. Then we calculate upon recalling (1.9) that for any P

$$\begin{aligned} A[\tilde{\mu}] + \bar{H}(P) &= \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(v,x) + H(Du,x)\, d\tilde{\mu} \\ &\geq \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} v\cdot(P+Dv)\, d\tilde{\mu} = P\cdot\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} v\, d\tilde{\mu} = P\cdot V. \end{aligned}$$

Consequently

$$A[\tilde{\mu}] \geq \max_P\{P\cdot V - \bar{H}(P)\} = \bar{L}(V).$$

By smoothing u as usual with a mollifer, we obtain this inequality in general. But in view of (5.13) we also have equality for the measure μ built above, and this establishes the identity (1.13). (The proof of (5.13) as written in [E3] invokes (1.13), but this fact is not really used.)

5.4 Useful formulas. We compute next the first and second derivatives of $\bar{H}^k$:

Theorem 5.3. *For $k = 1, \dots$ and $P \in \mathbb{R}^n$, we have the formulas*

$$D\bar{H}^k(P) = \int_{\mathbb{T}^n} D_pH(Du^k, x)\, d\sigma^k \tag{5.17}$$

and

$$\begin{aligned} D^2\bar{H}^k(P) = k\int_{\mathbb{T}^n} &(D_pH(Du^k, x)D^2_{xP}u^k - D\bar{H}^k(P)) \\ &\otimes (D_pH(Du^k, x)D^2_{xP}u^k - D\bar{H}^k(P))\, d\sigma^k \\ &+ \int_{\mathbb{T}^n} D^2_pH(Du^k, x)D^2_{xP}u^k \otimes D^2_{xP}u^k \; d\sigma^k. \end{aligned} \tag{5.18}$$

In particular, $\bar{H}^k$ is a convex function of P.

Proof. We have $e^{k\bar{H}^k(P)} = \int_{\mathbb{T}^n} e^{kH(Du^k,x)}dx$. Differentiate with respect to P_l:

$$\begin{aligned} ke^{k\bar{H}^k}\bar{H}^k_{P_l} &= \int_{\mathbb{T}^n} e^{kH(Du^k,x)}kH_{p_i}u^k_{P_lx_i}\, dx \\ &= \int_{\mathbb{T}^n} e^{kH(Du^k,x)}kH_{p_i}(\delta_{l,i} + v^k_{P_lx_i})\, dx = \int_{\mathbb{T}^n} e^{kH(Du^k,x)}kH_{p_l}\, dx, \end{aligned} \tag{5.19}$$

the last equality holding by (5.5). We cancel the k and rearrange, to derive (5.17).

As above, write out the l^{th} component of (5.17), and differentiate with respect to P_m:

$$\begin{aligned} e^{k\bar{H}^k}(\bar{H}^k_{P_lP_m} + k\bar{H}^k_{P_l}\bar{H}^k_{P_m}) = &\int_{\mathbb{T}^n} e^{kH(Du^k,x)}(H_{p_i}u^k_{P_mP_lx_i} + H_{p_ip_j}u^k_{P_lx_i}u^k_{P_mx_j})\, dx \\ &+ k\int_{\mathbb{T}^n} e^{kH(Du^k,x)}H_{p_i}u^k_{P_lx_i}H_{p_j}u^k_{P_mx_j}\, dx. \end{aligned}$$

The integral of the first term on the right is zero, and we can refashion the remaining expressions into formula (5.18). □

Application: nonresonance and averaging. Assume that $\bar{H}$ is differentiable at P and that $V = D\bar{H}(P)$ satisfies the *nonresonance condition*:

$$V \cdot m \neq 0 \quad \text{for each vector } m \in \mathbb{Z}^n, m \neq 0. \tag{5.20}$$

Theorem 5.4. *Suppose also that $D^2\bar{H}^k(P)$ is bounded as $k \to \infty$. Then*

$$\lim_{k\to\infty} \int_{\mathbb{T}^n} \Phi(D_Pu^k(P, x))\, d\sigma^k = \int_{\mathbb{T}^n} \Phi(X)\, dX \tag{5.21}$$

for each continuous, $\mathbb{T}^n$-periodic function Φ.

Proof. Observe first that the function $e^{2\pi i m\cdot D_P u^k} = e^{2\pi i m\cdot x}e^{2\pi i m\cdot D_P v^k}$ is $\mathbb{T}^n$-periodic. Hence we have

$$
\begin{aligned}
0 &= \int_{\mathbb{T}^n} D_p H(Du^k, x)\cdot D_x\left(e^{2\pi i m\cdot D_P u^k}\right) d\sigma^k \\
&= 2\pi i \int_{\mathbb{T}^n} e^{2\pi i m\cdot D_P u^k} m\cdot D_p H(Du^k, x) D^2_{xP} u^k \, d\sigma^k \\
&= 2\pi i \int_{\mathbb{T}^n} e^{2\pi i m\cdot D_P u^k} m\cdot D_P \bar{H}^k(P)\, d\sigma^k \\
&\qquad + 2\pi i \int_{\mathbb{T}^n} e^{2\pi i m\cdot D_P u^k} m\cdot (D_p H(Du^k, x) D^2_{xP} u^k - D_P \bar{H}^k(P))\, d\sigma^k.
\end{aligned}
$$

Consequently if $m \neq 0$, the identity (5.18) implies

$$
\begin{aligned}
|(m\cdot D_P\bar{H}^k(P))\int_{\mathbb{T}^n} e^{2\pi i m\cdot D_P u^k}\, d\sigma^k| &\leq |m|\int_{\mathbb{T}^n} |D_p H(Du^k, x) D^2_{xP} u^k - D_P\bar{H}^k(P))|\, d\sigma^k \\
&\leq |m|(k^{-1}\Delta \bar{H}^k(P))^{\frac{1}{2}}.
\end{aligned}
$$

Since $D\bar{H}^k(P) \to D\bar{H}(P) = V$ and $m\cdot V \neq 0$, we deduce that $\lim_{k\to\infty}\int_{\mathbb{T}^n} e^{2\pi i m\cdot D_P u^k}\, d\sigma^k = 0$, and so $\lim_{k\to\infty}\int_{\mathbb{T}^n} \Phi(D_P u^k(P, x))\, d\sigma^k = \int_{\mathbb{T}^n}\Phi(X)\, dX$ for each periodic function Φ whose Fourier expansion contains only finitely many nonzero terms. Such functions are dense in the sup–norm. □

Interpretation. Suppose that we regard $u = u(P, x)$ as a smooth generating function, inducing as in (1.5) the canonical change of variables $(p, x) \to (P, X)$, where $p = D_x u(P, x), X = D_P u(P, x)$. Then the corresponding Hamiltonian dynamics become (1.7) and consequently $\mathbf{X}(t) = Vt + X_0$. According to the nonresonance condition, we then have

$$
\lim_{T\to\infty}\frac{1}{T}\int_0^T \Phi(\mathbf{X}(t))\, dt = \int_{\mathbb{T}^n}\Phi(X)\, dX.
$$

Assertion (5.21) is consistent with this interpretation.

An important challenge for weak KAM theory is understanding further analytic consequences of (5.20), and stronger Diophantine conditions, beyond conventional application to small–divisor problems.

6. Quantum analogs.

This final section records a some observations and comments concerning the possible implications for quantum mechanics of weak KAM theory. I believe that investigating how to "quantize" weak KAM theory is definitely important, although it is not so clear whether these particular ideas will prove really useful. (A more honest title for this section would be "Quantum analogs ???".)

We will discuss only the Hamiltonian

$$H(p, x) := \frac{1}{2}|p|^2 + W(x), \tag{6.1}$$

in which the smooth potential W is $\mathbb{T}^n$-periodic. The corresponding stationary Schrödinger equation is

$$-\frac{h^2}{2}\Delta\psi + W\psi = E\psi \quad \text{in } \mathbb{R}^n, \tag{6.2}$$

E denoting the energy level and h Planck's constant.

6.1 Action minimizers. We propose as a quantum version of Mather's minimization principle to find a complex-valued state ψ minimizing the *action*

$$A[\psi] := \int_{\mathbb{T}^n} \frac{h^2}{2}|D\psi|^2 - W|\psi|^2 dx, \tag{6.3}$$

subject to the constraints that

$$\int_{\mathbb{T}^n} (\bar{\psi}D\psi - \psi D\bar{\psi}) \cdot D\phi \, dx = 0 \qquad \text{for all } \phi \in C^1(\mathbb{T}^n), \tag{6.4}$$

$$\int_{\mathbb{T}^n} |\psi|^2 \, dx = 1, \tag{6.5}$$

and

$$\frac{h}{2i} \int_{\mathbb{T}^n} \bar{\psi}D\psi - \psi D\bar{\psi} \, dx = V. \tag{6.6}$$

If ψ is smooth, condition (6.4) reads $\operatorname{div}(\mathbf{j}) = \operatorname{div}(\bar{\psi}D\psi - \psi D\bar{\psi}) = 0$: this is the analog of the flow invariance.

Let us take the complex-valued state in polar form

$$\psi = ae^{iu/h}, \tag{6.7}$$

where the phase $u = P \cdot x + v$ for some $\mathbb{T}^n$-periodic function v. Thus ψ has the *Bloch wave* form

$$\psi = e^{\frac{iP\cdot x}{h}}\phi, \tag{6.8}$$

for a periodic function ϕ. The action is then

$$A[\psi] = \int_{\mathbb{T}^n} \frac{h^2}{2}|Da|^2 + \frac{a^2}{2}|Du|^2 - Wa^2 \, dx, \tag{6.9}$$

and the constraints (6.4)–(6.6) become

$$\operatorname{div}(a^2 Du) = 0, \tag{6.10}$$

$$\int_{\mathbb{T}^n} a^2\, dx = 1, \tag{6.11}$$

$$\int_{\mathbb{T}^n} a^2 Du\, dx = V. \tag{6.12}$$

Quantum and classical action. We first present from [E2] an inequality between the effective Lagrangian $\bar{L}$ and the quantum action of a state ψ satisfying (6.4) – (6.6), for which $a^2 Du = \frac{h}{i} D\psi\, \bar{\psi}$ is periodic. To do so, we reintroduce our solution of the eikonal equation:

$$\frac{|D\hat{u}|^2}{2} + W = \bar{H}(P), \tag{C}$$

where $\hat{u} = P \cdot x + \hat{v}$ and $\hat{v}$ is periodic. To simplify notation later we are now introducing a circumflex for the solution of (C).

Theorem 6.1. *For each $P \in \mathbb{R}^n$ we have the equality*

$$A[\psi] - \bar{L}(V) + \bar{H}(P) + \bar{L}(V) - P \cdot V = \frac{h^2}{2} \int_{\mathbb{T}^n} |Da|^2 dx + \tfrac{1}{2} \int_{\mathbb{T}^n} |D\hat{u} - Du|^2\, a^2 dx. \tag{6.13}$$

In particular, if we take $P \in \partial \bar{L}(V)$, then

$$A[\psi] - \bar{L}(V) = \frac{h^2}{2} \int_{\mathbb{T}^n} |Da|^2\, dx + \tfrac{1}{2} \int_{\mathbb{T}^n} |D\hat{u} - Du|^2\, a^2 dx. \tag{6.14}$$

Proof. The left-hand side of (6.13) is

$$\begin{aligned} A[\psi] + \bar{H}(P) - P \cdot V = \int_{\mathbb{T}^n} & \frac{h^2}{2} |Da|^2 + \frac{|Du|^2}{2} a^2 - W a^2 dx \\ & + \int_{\mathbb{T}^n} \left(\frac{|D\hat{u}|^2}{2} + W\right) a^2 dx - \int_{\mathbb{T}^n} P \cdot Du\, a^2 dx. \end{aligned}$$

This equals

$$\frac{h^2}{2} \int_{\mathbb{T}^n} |Da|^2 dx + \tfrac{1}{2} \int_{\mathbb{T}^n} (|D\hat{u}|^2 + |Du|^2) a^2 dx - \int_{\mathbb{T}^n} (P + D\hat{v}) \cdot Du\, a^2 dx,$$

since $\hat{v}$, $a^2 Du$ are periodic and $\operatorname{div}(a^2 Du) = 0$. The foregoing simplifies to become the right hand side of (6.13). □

Observe that since $\bar{H}(P) + \bar{L}(V) \geq P \cdot V$, with equality if and only if $P \in \partial \bar{L}(V)$, this should be our optimal choice for P. Then (6.14) asserts that

$$\bar{L}(V) \leq A[\psi]$$

for all quantum states ψ satisfying (6.4) – (6.6), for the given average flux. So *the classical minimum of the action is a lower bound for the quantum action.* It would be extremely interesting to understand if and when this lower bound is approached as $h \to 0$.

The appendix of a recent paper by Gosse and Markowich [G-Mk] considers somewhat similar issues for the one-dimensional case.

First and second variation. ([E4]) We turn to a more careful analysis of our quantum action minimization problem. Let $\{(u(\tau), a(\tau))\}_{-1 \leq \tau \leq 1}$ be a smooth one-parameter family satisfying (6.10)–(6.12), with $(u(0), a(0)) = (u, a)$. Define

$$j(\tau) := \int_{\mathbb{T}^n} \frac{h^2}{2}|Da(\tau)|^2 + \frac{a^2(\tau)}{2}|Du(\tau)|^2 - Wa^2(\tau)\, dx$$

and write $' = \frac{d}{d\tau}$.

Theorem 6.2. (i) *We have $j'(0) = 0$ for all variations if and only if*

$$-\frac{h^2}{2}\Delta a = a\left(\frac{|Du|^2}{2} + W - E\right) \tag{6.15}$$

for some real number E.

(ii) *If* (6.15) *holds and $a > 0$, then*

$$j''(0) = \int_{\mathbb{T}^n} a^2|Du'|^2 + a^2\left|D(a'/a)\right|^2 dx.$$

Thus $j''(0) > 0$, provided $a' \neq 0$.

We interpret (6.15) as the *Euler–Lagrange equation* for our minimization problem.

Idea of proof. We will prove only assertion (i). First, compute

$$j' = \int_{\mathbb{T}^n} h^2 Da \cdot Da' + aa'|Du|^2 + a^2 Du \cdot Du' - 2Waa' dx.$$

Next, differentiate (6.10), (6.12):

$$\operatorname{div}(2aa'Du + a^2 Du') = 0 \tag{6.16}$$

$$\int_{\mathbb{T}^n} 2aa'Du + a^2 Du' dx = 0. \tag{6.17}$$

Recall that $Du = P + Dv$. Multiply (6.16) by v, integrate by parts, then take the inner product of (6.17) with P. Adding the resulting expressions gives us the identity $\int_{\mathbb{T}^n} 2aa'|Du|^2 + a^2 Du' \cdot Du\, dx = 0$. Hence

$$j' = \int_{\mathbb{T}^n} h^2 Da \cdot Da' - aa'|Du|^2 - 2Waa'dx = 2\int_{\mathbb{T}^n} a'\left(-\frac{h^2}{2}\Delta a - \left(\frac{|Du|^2}{2} + W\right) a\right) dx.$$

Then $j'(0) = 0$ for all such a', provided $-\frac{h^2}{2}\Delta a - \left(\frac{|Du|^2}{2} + W\right) a = -Ea$ for some real constant E. This is so since the variation a' must satisfy the identity $\int_{\mathbb{T}^n} a'a\, dx = 0$, had upon differentiating (6.11). □

6.2 Quasimodes. Our task now is constructing an explicit state ψ, which will turn out to be a critical point of $A[\cdot]$, subject to (6.4) – (6.6). Many of our computations are similar to those in the interesting paper of Anantharaman [A].

Dual eigenfunctions. We start with two linear problems. Consider the dual eigenvalue problems:

$$(6.18)\qquad \begin{cases} -\frac{h^2}{2}\Delta w + hP \cdot Dw - Ww = E^0 w & \text{in } \mathbb{T}^n \\ w \text{ is } \mathbb{T}^n\text{-periodic} \end{cases}$$

and

$$(6.19)\qquad \begin{cases} -\frac{h^2}{2}\Delta w^* - hP \cdot Dw^* - Ww^* = E^0 w^* & \text{in } \mathbb{T}^n \\ w^* \text{ is } \mathbb{T}^n\text{-periodic,} \end{cases}$$

where $E^0 = E^0(P) \in \mathbb{R}$ is the principal eigenvalue. We take the real eigenfunctions w, w^* to be positive in $\mathbb{T}^n$ and normalized so that $\int_{\mathbb{T}^n} ww^* dx = 1$. Define

$$v := -h \log w, \quad v^* := h \log w^*.$$

Then

$$(6.20)\qquad w = e^{-v/h}, \quad w^* = e^{v^*/h},$$

and a calculation shows that

$$(6.21)\qquad \begin{cases} -\frac{h}{2}\Delta v + \frac{1}{2}|P + Dv|^2 + W = \bar{H}_h(P) & \text{in } \mathbb{T}^n \\ v \text{ is } \mathbb{T}^n\text{-periodic,} \end{cases}$$

$$(6.22)\qquad \begin{cases} \frac{h}{2}\Delta v^* + \frac{1}{2}|P + Dv^*|^2 + W = \bar{H}_h(P) & \text{in } \mathbb{T}^n \\ v^* \text{ is } \mathbb{T}^n\text{-periodic,} \end{cases}$$

for

$$(6.23)\qquad \bar{H}_h(P) := \frac{|P|^2}{2} - E^0(P).$$

Transport and eikonal equations. Write

$$(6.24)\qquad \sigma := ww^* = e^{\frac{v^* - v}{h}}, \quad u := x \cdot P + \frac{v + v^*}{2}.$$

Then $\sigma > 0$ and $\int_{\mathbb{T}^n} \sigma\, dx = 1$.

Theorem 6.3. (i) *We have*

$$\operatorname{div}(\sigma Du) = 0 \quad \textit{in } \mathbb{T}^n. \tag{6.25}$$

Furthermore,

$$\frac{1}{2}|Du|^2 + W - \bar{H}_h(P) = \frac{h}{4}\Delta(v - v^*) - \frac{1}{8}|Dv - Dv^*|^2 \quad \textit{in } \mathbb{T}^n. \tag{6.26}$$

(ii) *In addition, these integral identities hold:*

$$\int_{\mathbb{T}^n} \frac{1}{2}|Du|^2 + W \, d\sigma = \bar{H}_h(P) + \frac{1}{8}\int_{\mathbb{T}^n} |Dv - Dv^*|^2 \, d\sigma \tag{6.27}$$

and

$$\int_{\mathbb{T}^n} |D^2u|^2 + \frac{1}{4}|D^2v - D^2v^*|^2 \, d\sigma = -\int_{\mathbb{T}^n} \Delta W \, d\sigma. \tag{6.28}$$

(iii) *Assume next that $\hat{u} = P \cdot x + \hat{v}$ solves (C). Then*

$$\frac{1}{2}\int_{\mathbb{T}^n} |\frac{1}{2}D(v + v^*) - D\hat{v}|^2 \, d\sigma + \frac{1}{8}\int_{\mathbb{T}^n} |Dv - Dv^*|^2 \, d\sigma = \bar{H}(P) - \bar{H}_h(P). \tag{6.29}$$

We call (6.25) the *transport equation*, and regard (6.26) as an *eikonal equation* with an error term.

It turns out that $\bar{H}(P) \to \bar{H}_h(P)$ as $h \to 0$. In addition, the measure $\sigma = \sigma_h$ converges weakly to a minimizing measure in the sense of §2. Anantharaman [A] provides much more interesting and detailed information about the limit measure, characterized in her paper as maximizing an entropy functional.

A quantum state. Now we define

$$\psi := ae^{\frac{iu}{h}}, \tag{6.30}$$

for

$$a := \sigma^{1/2} = e^{\frac{v^* - v}{2h}}. \tag{6.31}$$

We can calculate explicitly that

$$\frac{1}{2}|Du|^2 + W - \bar{H}_h(P) = -\frac{h^2}{2}\frac{\Delta a}{a} \quad \text{in } \mathbb{T}^n. \tag{6.32}$$

According then to Theorem 6.2, our ψ is a critical point of the action $A[\cdot]$. We can also introduce $\bar{L}_h$, the Legendre transform of $\bar{H}_h$, and write $V_h := D\bar{H}_h(P)$. It turns out that the quantum action of ψ is $A[\psi] = \bar{L}_h(V_h)$.

An approximate solution. We study next to what extent our ψ is an approximate solution of the stationary Schrödinger equation (6.2). Calculate:

$$
(6.33)\qquad \begin{aligned} -\frac{h^2}{2}\Delta\psi + W\psi - E\psi &= \left(\frac{1}{2}|Du|^2 + W - E\right)\psi - \frac{ih}{2}\frac{\operatorname{div}(a^2 Du)}{a^2}\psi - \frac{h^2}{2}\frac{\Delta a}{a}\psi \\ &=: A + B + C. \end{aligned}
$$

In view of (6.25), $B \equiv 0$. Now take $E = \bar{H}_h(P)$. According to (6.32), $A \equiv C$; that is, *the formal $O(1)$–term identically equals the formal $O(h^2)$–term in the expansion* (6.33). Therefore

$$
(6.34)\qquad -\frac{h^2}{2}\Delta\psi + W\psi - E\psi = 2(\frac{1}{2}|Du|^2 + W - \bar{H}_h(P))\psi
$$

for $E = \bar{H}_h(P)$.

Theorem 6.4. *If $E = \bar{H}_h(P)$,*

$$
(6.35)\qquad -\frac{h^2}{2}\Delta\psi + W\psi - E\psi = O(h),
$$

the right hand side estimated in $L^2(\mathbb{T}^n)$.

Idea of proof. Define the remainder term $R := 2(\frac{1}{2}|Du|^2 + W - \bar{H}_h(P))\psi$. Then

$$
\begin{aligned} \frac{1}{4}\int_{\mathbb{T}^n} |R|^2\,dx &= \int_{\mathbb{T}^n} \left(\frac{h}{4}\Delta(v - v^*) - \frac{1}{8}|D(v - v^*)|^2\right)^2 d\sigma \\ &= \int_{\mathbb{T}^n} \frac{h^2}{16}(\Delta(v - v^*))^2 - \frac{h}{16}\Delta(v - v^*)|D(v - v^*)|^2 + \frac{1}{64}|D(v - v^*)|^4\,d\sigma. \end{aligned}
$$

Integrating by parts in the middle term on the right leads after some simple estimates to the inequality

$$
(6.36)\qquad \int_{\mathbb{T}^n} |R|^2\,dx + \int_{\mathbb{T}^n} |D(v - v^*)|^4\,d\sigma \le Ch^2 \int_{\mathbb{T}^n} |D^2(v - v^*)|^2\,d\sigma.
$$

We then deduce from (6.28) that R is of order at most $O(h)$ in $L^2(\mathbb{T}^n)$. □

The final inequality shows that if $E = \bar{H}_h(P)$ and *if*

$$
(6.37)\qquad \int_{\mathbb{T}^n} |D^2(v - v^*)|^2 d\sigma = o(1),
$$

then $-\frac{h^2}{2}\Delta\psi + W\psi - E\psi = o(h)$ in $L^2(\mathbb{T}^n)$. We may hope that assertion (6.37) is true in some generality, although it can certainly fail, as has been shown by Y. Yu [Y]. *Further progress for this approach depends upon our indentifying conditions under which estimate*

(6.37) *holds*, and it is unclear to me what the prospects for this are. The nonresonance condition (5.20) will possibly be relevant for this.

Comparison with stochastic mechanics. I note in passing that there are striking formal connections with the Guerra–Morato and Nelson variational principle in stochastic quantum mechanics, as set forth in Nelson [N], Guerra–Morato [G-M], etc. See [E4] for some more discussion. I thank A. Majda for these references.

6.3 More homogenization. This section works out some relationships between $\bar{H}_h$ and homogenization theory for divergence–structure, second order elliptic PDE. (Cf. Bensoussan–Lions–Papanicolaou [B-L-P] and especially Capdeboscq [Cp]).

Let $A = ((a_{ij}))$ be symmetric, positive definite, and $\mathbb{T}^n$-periodic. Consider then this boundary value problem for an elliptic PDE with rapidly varying coefficients:

$$\begin{cases} -\left(a_{ij}\left(\frac{x}{\varepsilon}\right)u^\varepsilon_{x_i}\right)_{x_j} = f & \text{in } U \\ u^\varepsilon = 0 & \text{on } \partial U, \end{cases}$$

Then $u^\varepsilon \rightharpoonup u$ weakly in $H^1_0(U)$, u solving the limit problem

$$\begin{cases} -\bar{a}_{ij}u_{x_ix_j} = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases}$$

The *effective diffusion coefficient matrix* $\bar{A} = ((\bar{a}_{ij}))$ is determined this way. For $j = 1,\dots,n$, let χ^j solve the *corrector problem*

$$\begin{cases} -(a_{kl}\chi^j_{x_k})_{x_l} = (a_{jl})_{x_l} & \text{in } \mathbb{T}^n \\ \chi^j \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases} \tag{6.38}$$

Let us then for $i,j = 1,\dots,n$, set

$$\bar{a}_{ij} := \int_{\mathbb{T}^n} a_{ij} - a_{kl}\chi^i_{x_k}\chi^j_{x_l}\,dx. \tag{6.39}$$

Theorem 6.5. *Define* $V_h = D\bar{H}_h(P)$. *Then*

$$\bar{A}P = V_h, \tag{6.40}$$

where $\bar{A} = ((\bar{a}_{ij}))$ *is the effective coefficient matrix corresponding to* $A := a^2 I = ((a^2\delta_{ij}))$.

Proof. For the special case of the diagonal matrix A, the corrector PDE (6.38) reads

$$\begin{cases} -(a^2\chi^j_{x_k})_{x_k} = (a^2)_{x_j} & \text{in } \mathbb{T}^n \\ \chi^j \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases}$$

Now $u = x \cdot P + \frac{1}{2}(v + v^*)$ solves $\operatorname{div}(\sigma Du) = 0$, and therefore

$$-\operatorname{div}\left(a^2 \frac{1}{2} D(v + v^*)\right) = \operatorname{div}(a^2 P) = D(a^2) \cdot P.$$

Hence $\frac{v+v^*}{2} = P_i \chi^i$. Consequently for $j = 1, \dots, n$:

$$\begin{aligned} V_j &= \int_{\mathbb{T}^n} u_{x_j}\, d\sigma = \int_{\mathbb{T}^n} (\chi^i + x_i)_{x_j} P_i \, d\sigma = P_j - \int_{\mathbb{T}^n} \chi^i P_i \sigma_{x_j}\, dx \\ &= P_j + \int_{\mathbb{T}^n} \chi^i P_i (a^2 \chi^j_{x_k})_{x_k}\, dx = P_j - \int_{\mathbb{T}^n} P_i a^2 \chi^i_{x_k} \chi^j_{x_k}\, dx = (\bar{A}P)_j. \end{aligned}$$

□

Appendix: On notation.

This chart translates between some of our notation and that more customary in physics.

this paper	**meaning**	**typical physics notation**
u	action function	S
x	position	q
p	momentum	p
v	velocity	$\dot{q}$
X	'angle' variable	θ
P	action variable	I
V	'rotation' vector	ω
W	potential	V

This paper modifies some of the notation from earlier work, by now using "v, V" for variables previously denoted "q, Q" in the papers [E-G1-3, E2-3]. Our effective Hamiltonian $\bar{H}$ corresponds to Mather's function α and our $\bar{L}$ corresponds to his β.

References

[A-B] O. Alvarez and M. Bardi, *A general convergence result of singular perturbations of fully nonlinear degenerate parabolic PDEs*, to appear.

[A] N. Anantharaman, *Gibbs measures on path space and viscous approximation to action-minimizing measures*, Transactions AMS (2003).

[A-N] E. J. Anderson and P. Nash, *Linear Programming in Infinite Dimensional Spaces*, Wiley, 1987.

[A-K-N] W. I. Arnold, W. W. Kozlov and A. I. Neishtadt, *Mathematical aspects of classical and celestial mechanics*, Encyclopedia of Mathematical Sciences: Dynamical Systems III, Springer, 1988.

[Au] S. Aubry, *The twist map, the extended Frenkel–Kantorova model and the devil's straircase*, Physica D **7** (1983), 240–258.

[B-S] G. Barles and P. Souganidis, *Some counterexamples on the asymptotic behavior of the solutions of Hamilton–Jacobi equations*, C. R. Acad. Sci. Paris Sr. I Math. **330** (2000), 963–968.

[B] E. N. Barron, *Viscosity solutions and analysis in L^∞*, in Nonlinear Analysis, Differential Equations and Control (1999), Dordrecht, 1–60.

[B-L-P] A. Bensoussan, J.-L. Lions and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North-Holland, 1978.

[Be] P. Bernard, *The action spectrum near positive definite invariant tori*, to appear.

[B-T] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, 1997.

[B-K] A. Bourlioux and B. Khouider, *Computing the effective Hamiltonian in the Majda–Souganidis model of turbulent premixed flames*, SIAM J Numerical Analysis **40** (2002), 1330–1353.

[Cp] Y. Capdeboscq, *Homogenization of a neutronic critical diffusion problem with drift*, Proc. Royal Soc Edinburgh **132** (2002), 567–594.

[C1] M. Concordel, *Periodic homogenization of Hamilton-Jacobi equations I: additive eigenvalues and variational formula*, Indiana Univ. Math. J. **45** (1996), 1095–1117.

[C2] M. Concordel, *Periodic homogenisation of Hamilton-Jacobi equations II: eikonal equations*, Proc. Roy. Soc. Edinburgh **127** (1997), 665–689.

[C-I] G. Contreras and R. Iturriaga, *Global minimizers of autonomous Lagrangians*, Lecture notes.

[D] P. Deift, *Integrable Hamiltonian systems*, Lectures in Applied Mathematics **31** (1996), 103–138.

[DC] M. J. Dias Carneiro, *On minimizing measure of the action of autonomous Lagrangians*, Nonlinearity **8** (1995), 1077–1085.

[EW] Weinan E, *Aubry–Mather theory and periodic solutions of the forced Burgers equation*, Comm Pure and Appl Math **52** (1999), 811–828.

[E1] L. C. Evans, *Periodic homogenization of certain fully nonlinear PDE*, Proc Royal Society Edinburgh **120** (1992), 245–265.

[E2] L. C. Evans, *Effective Hamiltonians and quantum states*, Seminaire Equations aux Dérivées Partielles, Ecole Polytechnique (2000–2001).

[E3] L. C. Evans, *Some new PDE methods for weak KAM theory*, Calculus of Variations and PDE **17** (2003), 159–177.

[E4] L. C. Evans, *Towards a quantum analog of weak KAM theory*, to appear in Comm. Math Physics.

[E-G1] L. C. Evans and D. Gomes, *Effective Hamiltonians and averaging for Hamiltonian dynamics I*, Archive Rational Mech and Analysis **157** (2001), 1–33.

[E-G2] L. C. Evans and D. Gomes, *Effective Hamiltonians and averaging for Hamiltonian dynamics II*, Archive Rational Mech and Analysis **161** (2002), 271–305.

[E-G3] L. C. Evans and D. Gomes, *Linear programming interpretations of Mather's variational principle*, ESIAM: Control, Optimisation and Calculus of Variations **8** (2002).

[F1] A. Fathi, *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*, C. R. Acad. Sci. Paris Sr. I Math. **324** (1997), 1043–1046.

[F2] A. Fathi, *Solutions KAM faibles conjuguées et barrières de Peierls*, C. R. Acad. Sci. Paris Sr. I Math. **325** (1997), 649–652.

[F3] A. Fathi, *Orbites hétéroclines et ensemble de Peierls*, C. R. Acad. Sci. Paris Sr. I Math. **326** (1998), 1213–1216.

[F4] A. Fathi, *Sur la convergence du semi-groupe de Lax-Oleinik*, C. R. Acad. Sci. Paris Sr. I Math. **327** (1998), 267–270.

[F5] A. Fathi, *Weak KAM theory in Lagrangian Dynamics, Preliminary Version*, Lecture notes (2003).

[Fa-M] A. Fathi and J. Mather, *Failure of convergence of the Lax–Oleinik semigroup in the time–periodic case*, Bull. Soc. Math. France **128** (2000), 473–483.

[F-Ma] A. Fathi and E. Maderna, *Weak KAM theorem on noncompact manifolds*, to appear.

[F-S1] A. Fathi and A. Siconolfi, *Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, to appear in Invent. Math.

[F-S2] A. Fathi and A. Siconolfi, *PDE aspects of Aubry–Mather theory for continuous convex Hamiltonians*, to appear.

[Fo-M] G. Forni and J. Mather, *Action minimizing orbits in Hamiltonian systems*, Transition to Chaos in Classical and Quantum Mechanics, Lecture Notes in Math 1589 (S. Graffi, ed.), Sringer, 1994.

[Gd] H. Goldstein, *Classical mechanics* (2nd ed.), Addison-Wesley, 1980.

[G1] D. Gomes, *Hamilton–Jacobi Equations, Viscosity Solutions and Asymptotics of Hamiltonian Systems*, Ph.D. Thesis, University of California, Berkeley (2000).

[G2] D. Gomes, *Viscosity solutions of Hamilton-Jacobi equations and asymptotics for Hamiltonian systems*, Calculus of Variations and PDE **14** (2002), 345–357.

[G3] D. Gomes, *Perturbation theory for Hamilton-Jacobi equations and stability of Aubry-Mather sets*, to appear in SIAM journal of Mathematical Analysis.

[G4] D. Gomes, *A stochastic analog of Aubry-Mather theory*, Nonlinearity **10** (2002), 271-305.

[G5] D. Gomes, *Regularity theory for Hamilton-Jacobi equations*, J. Differential Equations **187** (2003), 359–374.

[G-O] D. Gomes and A. Oberman, *Computing the effective Hamiltonian: a variational approach to homogenization*, to appear.

[G-Mk] L. Gosse and P. Markowich, *Multiphase semiclassical approximation of an electron in a one–dimensional crystalline lattice. I: Homogeneous problems, vanishing exterior potentials*, to appear.

[G-M] F. Guerra and L. Morato, *Quantization of dynamical systems and stochastic control theory*, Physical Review D **27** (1983), 1774–1786.

[J-K-M] H. R. Jauslin, H. O. Kreiss and J. Moser, *On the forced Burgers equation with periodic boundary conditions*, in Proc. Sympos. Pure Math **65** (1999), 133–153.

[L] P. Lax, *Linear Algebra*, John Wiley, 1997.

[L-P-V] P.-L. Lions, G. Papanicolaou, and S. R. S. Varadhan, *Homogenization of Hamilton–Jacobi equations*, unpublished, circa 1988.

[L-S] P. L. Lions and P. Souganidis, *Corrector for the homogenization of Hamilton–Jacobi equations in the stationary ergodic setting*, Comm Pure Applied Math **56** (2003), 1501–1524.

[M-S] A. Majda and P. Souganidis, *Large scale front dynamics for turbulent reaction–diffusion equations with separated velocity scales*, Nonlinearity **7** (1994), 1–30.

[Mn1] R. Mañé, *Global Variational Methods in Conservative Dynamics*, Instituto de Matemática Pura e Aplicada, Rio de Janeiro.

[Mn2] R. Mañé, *On the minimizing measures of Lagrangian dynamical systems*, Nonlinearity **5** (1992), 623–638.

[Mn3] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (1996), 273–310.

[Mt1] J. Mather, *Minimal measures*, Comment. Math Helvetici **64** (1989), 375–394.

[Mt2] J. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Zeitschrift **207** (1991), 169–207.

[Mt3] J. Mather, *Differentiability of the minimal average action as a function of the rotation number*, Bol. Soc. Bras. Math **21** (1990), 59–70.

[Mt4] J. Mather, *Variational construction of connecting orbits*, Ann. Inst. Fourier, Grenoble **43** (1993), 1347–1386.

[N] E. Nelson, *Quantum Fluctuations*, Princeton University Press, 1985.

[O] A. Oberman, *Some numerical experiments with homogenization*, preprint (2003).

[Q] J. Qian, *Two approximations for effective Hamiltonians arising from homogenization of Hamilton–Jacobi equation*, preprint (2003).

[Y] Y. Yu, *Error estimates for a quantization of a Mather set*, preprint (2002).